

## **ЗАСТОСУВАННЯ МАРКОВСЬКИХ МОДЕЛЕЙ ІЗ ВІРОГІДНІСТЮ ДЛЯ МАШИННОГО ПЕРЕКЛАДУ**

**І.М. СИВОПЛЯС<sup>1\*</sup>, О.В. ФІЛОНЕНКО<sup>1</sup>, О.П. ЧЕРНИХ<sup>2</sup>, О.М. ШЕІН<sup>2</sup>**

<sup>1</sup> студент кафедри ОТП, НТУ «ХПІ», Харків, УКРАЇНА

<sup>2</sup> доцент кафедри ОТП, НТУ «ХПІ», Харків, УКРАЇНА

\* email: ivan.kozaks@gmail.com

Машинний переклад – технологія, що дозволяє отримати переклад із мови оригіналу у мову перекладання за допомогою комп'ютера. У теперішній час сучасний машинний переклад дозволяє отримати достатньо коректне розуміння тексту оригіналу, але порівняно із людським перекладом має два суттєвих недоліки:

- інтерпретація подвійного змісту слів (так звана акуратність);
- неправильні спрацьовування (так зване покриття).

Найбільш поширенішими методами позбавлення від цих недоліків є використання моделей на основі правил та моделей на основі статистики та теорії ймовірностей. Моделі на основі правил – моделі, що використовують правила для перекладу з однієї мови на іншу. Моделі на основі статистики – моделі, що базуються на статистиці використання слів у текстах. Марковські моделі засновані на статистиці та теорії ймовірностей і дозволяють суттєво скоротити тренування та правильне визначення нових слів. Вони покращують "розуміння" системою контексту слів, визначення залежностей, добування інформації з тексту, пошук та безліч інших можливостей. Із існуючих недоліків можна виділити порівняно великі об'єми форматованих текстів (корпусів), які необхідні для навчання та тренування таких моделей.

На теперішній час є багато пропозицій по вирішенню цих проблем. Для визначення оптимальної моделі використовуються три критерії відбору моделі для машинного перекладу:

- правильність інтерпретації слів із подвійним змістом;
- мінімальна кількість неправильних спрацьовувань;
- пристосованість до появи нових слів.

Найкращим рішенням, що задовольняє цим критеріям, буде гібридна модель із використанням Марковських моделей для загального перекладу та додаткових правил з метою поліпшення машинного перекладу і позбавлення слів подвійного змісту. Таким чином, можна досягти майже стовідсотково правильного перекладу тексту в залежності від налаштування додаткових правил та ширини лексики тренувального корпусу, що дасть можливість ще на один крок наблизитися до ідеального машинного перекладу.